

Article

## Distributed Pedestrian Detection Alerts Based on Data Fusion with Accurate Localization

Fernando García <sup>1,\*</sup>, Felipe Jiménez <sup>2</sup>, José Javier Anaya <sup>2</sup>, José María Armingol <sup>1</sup>,  
José Eugenio Naranjo <sup>2</sup> and Arturo de la Escalera <sup>1</sup>

<sup>1</sup> Intelligent Systems Laboratory, Universidad Carlos III de Madrid, Avda de la Universidad 30, Leganés 28911, Madrid, Spain; E-Mails: armingol@ing.uc3m.es (J.M.A.); escalera@ing.uc3m.es (A.D.L.E.)

<sup>2</sup> INSIA, Technical University of Madrid, Carretera de Valencia, km.7, Madrid 28031, Spain; E-Mails: felipe.jimenez@upm.es (F.J.); jj.anaya@upm.es (J.J.A.); joseeugenio.naranjo@upm.es (J.E.N.)

\* Author to whom correspondence should be addressed; E-Mail: fegarcia@ing.uc3m.es; Tel.: +34-916-248-813; Fax: +34-916-249-430.

Received: 5 July 2013; in revised form: 27 August 2013 / Accepted: 30 August 2013 /  
Published: 4 September 2013

---

**Abstract:** Among Advanced Driver Assistance Systems (ADAS) pedestrian detection is a common issue due to the vulnerability of pedestrians in the event of accidents. In the present work, a novel approach for pedestrian detection based on data fusion is presented. Data fusion helps to overcome the limitations inherent to each detection system (computer vision and laser scanner) and provides accurate and trustable tracking of any pedestrian movement. The application is complemented by an efficient communication protocol, able to alert vehicles in the surroundings by a fast and reliable communication. The combination of a powerful location, based on a GPS with inertial measurement, and accurate obstacle localization based on data fusion has allowed locating the detected pedestrians with high accuracy. Tests proved the viability of the detection system and the efficiency of the communication, even at long distances. By the use of the alert communication, dangerous situations such as occlusions or misdetections can be avoided.

**Keywords:** Advance Driver Assistance Systems; VANETs; V2V; data fusion; laser scanner and computer vision

---

## 1. Introduction

Of all the problems that are related to transportation, traffic accidents are the most dramatic since they deal with human lives. The efforts of recent years, such as an increase in the safety measures for roads and vehicles or the enhancement of traffic laws to decrease drivers' misbehaviors, have led to reduced death tolls in road accidents, yet each year in the European Union more than 1 million road accidents still occur in which over 31 thousand people die [1]. Thus, even though the efforts made are helping to mitigate this number, there is still a considerable amount of work to be done. The new information and communication technologies developed in the last decade have enabled more complex and reliable safety applications to be created. These new applications are able to reduce the number of accidents and deaths on the road by both preventing them and abating the harm caused by accidents. Most traffic accidents are related to human errors. Carelessness and erroneous decisions by the driver are the two main factors that cause traffic accidents. These kinds of errors, related with human nature, are impossible to be eliminated, although efforts can be made to decrease them. Recent research in the field of Intelligent Vehicles (IV) has focused on using advances in information technologies to prevent these errors. Advanced Driver Assistance Systems (ADAS) try to warn the driver and prepare the driver in the event of hazardous situations. During recent years, Intelligent Transportation Systems (ITS) research projects have been abundant in Spain, with several laboratories working in different fields related to the topic at hand. INSIA and Universidad Complutense de Madrid have presented different works related to the time to collision calculation for vehicle intersections [2] and autonomous maneuvers for vehicle avoidance [3]. On the other hand, Carlos III University has proposed several works related with sensor detection classification and tracking [4]. In recent years, these works, sponsored by the Spanish government, have led to the presentation of several works on the topic [5,6]. Other laboratories in Spain are also working on the topic with important contributions, such as the works presented by Universitat Autònoma de Barcelona [7,8] and Universidad de Alcalá de Henares [9,10].

The work presented in this paper represents a step forward in pedestrian detection and danger communication. The use of Data Fusion techniques to enhance the capabilities of classic Advanced Driver Assistance Systems allows improving the performance of the detection system, based on the use of two well-known solutions, laser scanners and computer vision. Furthermore, the communication system allows providing accurate and fast pedestrian detection based on the precise location of the vehicle and the use of a very accurate obstacle detection device, *i.e.*, a laser scanner. The system is based on the occlusion scenario, where pedestrians are occluded by other vehicles on the road. This situation is frequent in urban environments and represents one of the most typical accidents involving pedestrians. The proposed solution consists of the fact that the vehicle that occludes a pedestrian automatically informs other vehicles in the surroundings of the presence of the pedestrian by means of a pedestrian detection and alert communication protocol. In these situations, the availability of accurate sensors, able to detect pedestrians with robustness, is as important as the capability to inform the vehicles in the surroundings, based on trustable technology. The communication to the vehicles has to be fast and reliable and able to adapt to fast changing scenarios such as urban environments.

Among the different contributions presented in this paper, we can highlight an improved laser scanner-based pedestrian detection system, the high level fusion system that combines information

from a laser scanner and artificial vision, together with the tracking system, and the novel track definition (consolidated/non-consolidated).

## 2. Related Work

Data Fusion is becoming frequent in ITS works, due to the need for reliable and trustable sensor systems for road safety applications. The fusion of different sensing devices helps to overcome the limitation of each sensing technology, providing accurate and trustable detections.

Fusion approaches in vehicle safety are divided according to the fusion architecture used. Most of these works focus on the classification process:

In *decentralized* approaches, detections and classifications are performed by an independent subsystem with limited information (typically a single sensor). A final stage combines the information and classification, according to the sensors and the certainty of the detections. In [11], Adaboost vision-based pedestrian detection and Gaussian Mixture Model classifier (GMM) for laser scanner-based pedestrian detections, a Bayesian decisor is used to combine detections of both subsystems. In [12], multidimensional features for laser scanner pedestrian detection are used, and Histograms of Oriented Gradients (HOG) features and Support Vector Matching (SVM) for visual detection is performed by a Bayesian model approach. In [13], a similar approach is presented in comparison with other medium level algorithms.

In *centralized* architectures, a set of combined features from different sensors is created. Hence a single classification stage is necessary with the information from all the available sensors. A classical centralized data fusion approach for pedestrian detection is stereo vision systems, although some authors tend to consider stereo vision as a single scanner. Other works take advantage of different sensing technologies: In [13,14], a complete work with tests of different algorithms is presented to combine the features from different sensing devices, again based on laser scanner and computer vision. The alternative methods presented are Naïve Bayes, GMMC, NN, FLDA, and SVM. The work in [15] uses feature level information to improve the tracking of the objects.

Other approaches take advantage of the trustability of the laser scanner to detect the Region of Interest (ROI) in an image and perform computer vision detection taking advantage of the amount of information available from the computer vision sensor. In [16] SVM is used for vision-based vehicle detection and classification, while [17] uses Convolutional Neural Networks for pedestrian detection and [18] HOG and SVM approach. Finally, [19] uses Invariant Features and SVM to perform vision-based pedestrian detections. Based on a different point of view, the authors of [20] search in the environment for dangerous zones where pedestrians could be located, based on the laser scanner information (e.g., space between two vehicles). In those regions, pedestrian classification is performed based on a vision approach. Although these approaches deal with several sensors, each process (classification and detection) is performed by a single sensor, thus the fusion process is limited to the data alignment process.

It is clear that the information retrieved by a vehicle provided only by its local sensors should be enough to prevent near accidents or to reduce the effects of a certain accident. However, new vehicle applications such as Cooperative Collision Warning [21], Cross-Flow Turn Assistant [22], Curve Speed Warning [23], Emergency Vehicle Warning [24] and others, require additional information on

the circulation environment, from the other vehicles as well as from the infrastructure. This means that an external source of information is necessary in the vehicle itself in order to provide the necessary information to guarantee a proper performance of these assistance systems. Vehicle to Vehicle (V2V) and Vehicle to Infrastructure (Roadside) (V2I) communications cover the gap of providing information to the Advanced Driver Assistance Systems (ADAS) from the surroundings of a vehicle.

### 3. General Description

When dealing with safety applications, a fast and reliable obstacle identification and report method is required. But perception sensors have limitations due to the technology used (e.g., range limitation, usability, occlusions). The combination of different sensors can overcome these problems. But urban scenarios are fast-changing and unstructured. Thus, situations arise where detections are unfeasible. A distributed pedestrian detection can help to detect a pedestrian even in the most challenging scenarios, allowing the vehicle to perceive information that would be unattainable in a normal situation. This distributed information can be provided by other vehicles or the infrastructure.

The work at hand focuses on a distributed detection and alert of pedestrians. A fusion approach is used for reliable pedestrian detection and localization, and a trustable communication protocol alerts other vehicles of the situation in advance. In the fusion architecture created, the Kalman Filter (KF) approach is included for pedestrian tracking. The Kalman Filter represents an important part of the approach allowing us to track pedestrians and to fuse the information of the detection of both sensors, as explained in Section 4.

Two vehicles were used (Figure 1). The first one is a research platform for ADAS test and development, with multiple sensing devices incorporated. The second platform is a vehicle equipped with an on-board computer able to display the detection to the driver. In the present application three sensing devices were mainly used, an accurate high definition GPS with inertial measurement, used for accurate localization of the vehicle and the pedestrians, and two sensors used for pedestrian detection (laser scanner and computer vision). Both vehicles were equipped with the communication devices (MTM-CM3100 gateways, Maxfor Inc., Santa Clara, CA, USA) that provided the communication channel vehicles.

**Figure 1.** Left platform IVVI 2.0 with all the sensing devices available. Right vehicle with the communication receiver.



The work represents a step forward, combining a state-of-the-art pedestrian detection algorithm with the capacity of Vehicle to Vehicle communications (V2V), creating a robust pedestrian detection and alert communication.

#### 4. Data Fusion Based Pedestrian Detection

As previously mentioned, the pedestrian detection approach presented is based on the use of two well-known technologies in Intelligent Transportation Systems, a laser scanner and computer vision. The former is a recent sensor in automotive applications that helps in the process of creating intelligent applications thanks to its trustworthiness and reliability, but the information provided is limited. On the other hand computer vision provides a high amount of information. This information is unstructured, and with limited reliability. The combination of both systems helps to overcome the limitations inherent to each system resulting in a reliable application that fulfills the key requirements of safety applications.

In this section, a pedestrian detection system is described for each subsystem (laser scanner and computer vision). Also, procedures to extrapolate laser scanner detections to the computer vision coordinate system are specified; finally, all these detections must be extrapolated to the vehicle coordination system that is the front part of the vehicle. Using a high precision GPS with an inertial measurement system (MTI-G) from Xsens Technologies B.V. (Enschede, The Netherlands), it is possible to obtain precise localization of the vehicle and thus of the pedestrians around. All this information is provided to the second vehicle, obtaining accurate localization of the pedestrians detected.

##### 4.1. Laser Scanner Pedestrian Detection

The laser scanner was mounted in the bumper of the test platform IVVI 2.0 (Figure 1). Before pedestrian classification, obstacle segmentation is mandatory, with shape estimation. Later, this information is used to select which of the detected obstacles are suitable to be classified as pedestrians. Later, obstacle classification is performed taking into account the previously estimated shape.

##### Obstacle Segmentation and Shape Estimation

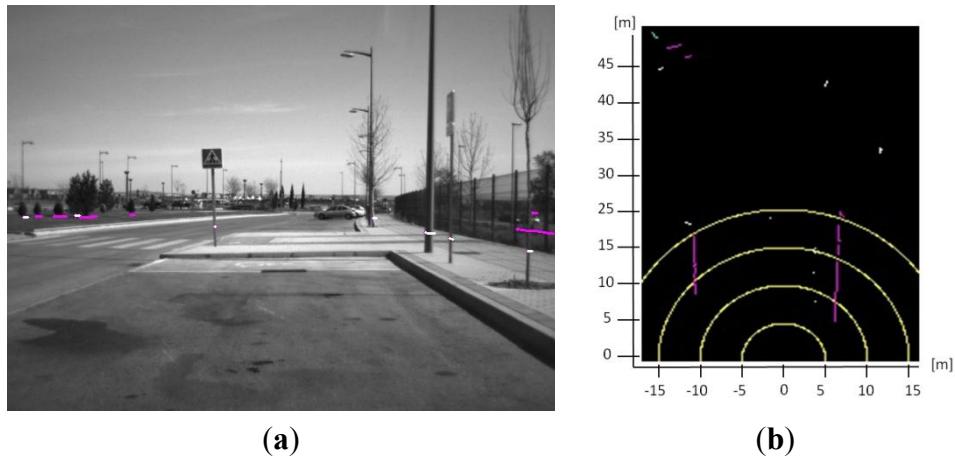
Laser scanner rotation provides 401 distance points per scan according to the rotation angle with 0.25 degrees of resolution, each one of them is provided with a time delay with respect to the others. Thus, movement of the vehicle has to be compensated using the information given by the high precision GPS system. Euler angles displacement and velocity must be taken into account in order to provide accurate shape reconstruction. Furthermore, the laser scanner is also very sensitive to pitching movements, so this information is also used to check if there is a strong pitch movement that can lead to misdetections. Applying the compensation with the rotation and translation matrixes (Equation (1)) the points are referenced to the position of the last point received (Figure 2).

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = R \begin{bmatrix} x_0 \\ y_0 \\ z_0 \end{bmatrix} + T + T_0$$

$$R = \begin{bmatrix} \cos(\Delta\delta) & 0 & \sin(\Delta\delta) \\ 0 & 1 & 0 \\ -\sin(\Delta\delta) & 0 & \cos(\Delta\delta) \end{bmatrix} \begin{bmatrix} 1 & 0 & 1 \\ 0 & \cos(\Delta\varphi) & -\sin(\Delta\varphi) \\ 0 & \sin(\Delta\varphi) & \cos(\Delta\varphi) \end{bmatrix} \begin{bmatrix} \cos(\Delta\theta) & -\sin(\Delta\theta) & 0 \\ \sin(\Delta\theta) & \cos(\Delta\theta) & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad (1)$$

$$T_0 = \begin{bmatrix} vT_i \times \cos(\Delta\varphi) \\ vT_i \times \sin(\Delta\varphi) \\ 0 \end{bmatrix}, T_0 = \begin{bmatrix} x_t \\ y_t \\ z_t \end{bmatrix}$$

**Figure 2.** Laser scanner clustering and shape reconstruction. **(a)** The points are shown in the image field; **(b)** the information only from the laser scanner is shown with the polylines reconstruction, where yellow lines represent the distances 5, 10, 15 and 20 m to the laser scanner.



In Equation (1),  $\Delta\delta$ ,  $\Delta\varphi$  and  $\Delta\theta$  correspond to the increment of the Euler angles roll, pitch and yaw respectively for a given period of time  $T_i$ . Coordinates  $(x,y,z)$  and  $(x_0,y_0,z_0)$  are the Cartesian coordinates of a given point after the vehicle movement compensation.  $R$  is the rotation matrix,  $T_v$  the translation matrix according to the velocity of the vehicle,  $T_0$  the translation matrix according to the position of the laser and the inertial sensor. Finally, after clustering the shapes of the different obstacles are estimated using polylines. Polyline creation as well as detailed explanation of the clustering algorithm is given in [25] (Figure 2).

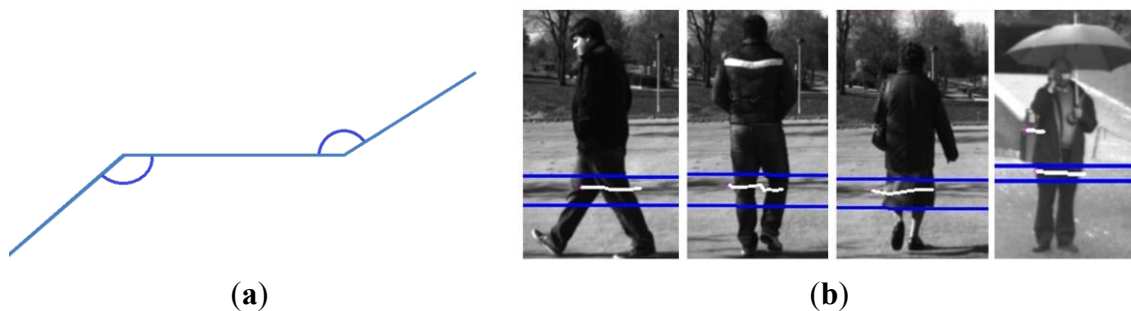
### Obstacle Classification

After shape estimation, classification is performed, differentiating among the following kind of obstacles: Big obstacles, Road limits, Vehicles, L shaped, Small obstacles and Pedestrians. Present applications focus on pedestrian detection, and a detailed explanation of the other types of obstacles is provided in [6].

Final pedestrian classification is performed in three steps, using *a priori* contextual information. Firstly, obstacles with a size proportional to a pedestrian are selected, based on anthropometric studies

among the different obstacles; later, the shape of the polyline is checked with a typical pedestrian pattern; finally, a tracking stage is included, using the anthropometric information to avoid false positives. The idea behind this method was presented for the first time in [25], although the high rate of false positives obtained required a redesign of certain aspects, *i.e.*, tracking stage to avoid false positives or the use of context information based on anthropometric studies. A study of the different patterns given by pedestrians was performed giving the pattern shown in Figure 3a.

**Figure 3.** (a) Pattern for pedestrian detection. (b) Different examples of different patterns given by pedestrians with different leg positions. The blue lines represent the boxes including pedestrian detections of the laser scanner.



### Pattern Matching

In this pattern three polylines are presented and the angles that connect the polylines are included within the limits of  $[0, \frac{\pi}{2}]$ .

Pattern matching process computes the two angles and gives a similarity score:

$$Similarity = \frac{2\theta_1}{\pi} \times \frac{2\theta_2}{\pi} \quad (2)$$

where  $\theta_1$  and  $\theta_2$  are the angles that connect two consecutive lines.

This similarity is computed among consecutive polylines that fit with the pedestrians' size. If the result is greater than a given threshold, the obstacle is considered to be a pedestrian. It is assumed that a previous pattern is common in unstructured scenarios, thus false positives are expected. Using information fusion with computer vision, this limitation can be overcome.

The specified pattern proved to be a useful tool, even in scenarios with certain difficulties, such as pedestrians wearing skirts or static pedestrians, where the detection of pedestrians using the specified pattern represents a higher challenge. As shown in Figure 3b, the results provided by the pattern in these specific scenarios were very positive. The similarity threshold was empirically chosen after the study of several sequences in controlled scenarios, selecting the threshold that assured positive rates over 80% in all movements and static pedestrians.

Information provided by the laser scanner is limited, due to the specific nature of the sensor, to a 401 distance. Thus, providing a trustable classification is a relatively difficult task. In order to avoid misdetection, a higher level stage was designed to avoid errors in the classification process. This stage consisted of the estimation of the movement based on KF, and feature-based correlation (features). This tracking stage helps to study the movement of the pedestrians, eliminating those pedestrians that



do not fit with the movement of a pedestrian (e.g., fast speeds or size changes). Finally, there is an upper level classification based on a voting scheme, according to the classification in the last ten scans:

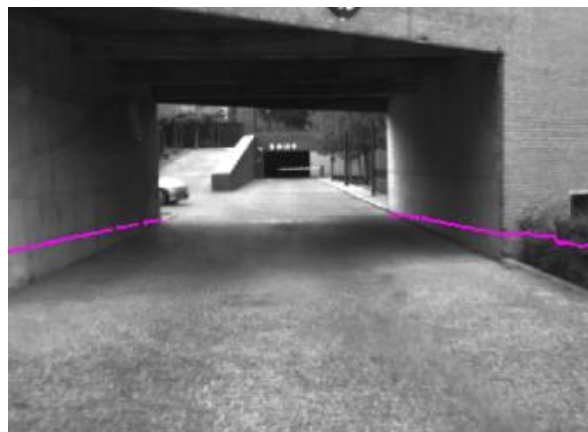
$$\text{Features: } < \bar{N}, \text{height}, \text{width}, \delta_{+x}, \delta_{-x}, \sigma, d, \rho >$$

Obstacle features used for comparison, where  $\bar{N}$  is the medium number of points, height and width are the sizes of the obstacle,  $\sigma$  is the standard deviation of the points to the center of the obstacle.  $\rho$  is the radius of the circle that surrounds the obstacle and  $d$  is the distance to the estimation of the KF. Finally  $\delta_{+x}$ ,  $\delta_{-x}$  are the numbers of points to the left or to the right of the center. Before ROI detection some data alignment must be performed since sensors do not share the same coordination system. This data alignment is explained in the following section.

#### 4.2. Vision Based Pedestrian Detection

Laser scanner detections are extrapolated to the field of view of the camera, thus only obstacles given by the laser scanner and extrapolated to the image are processed to check whether they are pedestrians or not. This way, the reliability of the laser scanner is used to reduce the false positives provided by the vision systems. In addition, the amount of information used for visual processing is reduced by reducing the region of the images processed by the computer vision system to those provided by the laser scanner (Figure 4), allowing real time processing. If the laser scanner is unavailable (e.g., strong pitch movement), the whole image is used to process the pedestrian detection.

**Figure 4.** Laser scanner detection extrapolated to the camera field of view with accurate extrinsic calibration.



To provide the laser scanner information to the camera coordination system a pin-hole model was used together with a high-accuracy extrinsic calibration system. Both coordination frames of the two different subsystems were translated to the reference point, which is the central point of the front bumper of the vehicle. The calibration process was performed online, allowing recalibration in real time of the Euler angles in case of necessity. The calibration process is performed online and supervised. Once the equipment is mounted and the laser scanner detection points are displayed on the image, they are varied online until they match with the real image. Supervised online calibration is a process performed taking into account the extreme sensitivity of the laser scanner to pitch angle.

To perform this coordinate change, transformation and rotation matrix must be used:



$$\begin{bmatrix} x_v \\ y_v \\ z_v \end{bmatrix} = R_i \left( \begin{bmatrix} x_i \\ y_i \\ z_i \end{bmatrix} \right) + T_i \quad (3)$$

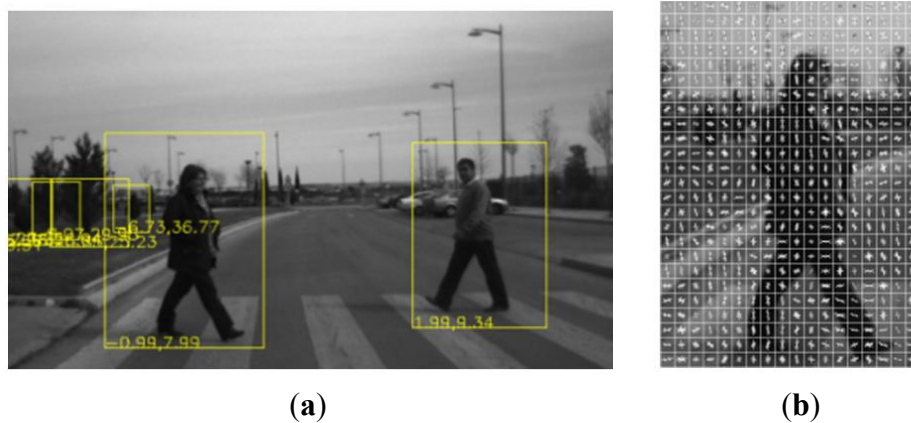
where  $T_i$  is the translation matrix and  $R_i$  is the rotation. This rotation matrix is equivalent to the rotation matrix used in Equation (1). but in this case, Euler angles correspond to the angular deviation among the coordinate system.  $T_i$  is equivalent to the displacement in the Cartesian coordinate system for each sensor:

$$\lambda \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} f & 0 & u_0 \\ 0 & f & v_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_c \\ z_c \\ y_c \end{bmatrix} \quad (4)$$

where  $u, v$  are the image coordinates in pixels  $f$  the focal distance and  $(x_c, y_c, z_c)$  the Cartesian coordinates of the image detections in the image coordinate system.  $u_0$  and  $v_0$  are the coordinates of the center of the image.

Equations (3) and (4) were also used to transform laser scanner obstacles that fit with pedestrian size to camera coordinate system. Providing Regions of Interest (ROIs) where the classification is performed. This way, obstacle association from both sensors is implicit, since they perform classification for each sensor over the same set of obstacles provided by the laser scanner. This way accuracy and trustability of the laser scanner is used to provide trustable and accurate obstacle detection (Figure 5).

**Figure 5.** Laser scanner ROI detection and HOG features computed in one of the regions. Figure (a) shows the boxes that represent the ROIs given by the laser scanner with the distances highlighted; (b) shows the computation of the HOG features in one of these ROIs.



Vision based pedestrian detection is based on the Histogram of Oriented Gradients (HOG) descriptor and Support Vector Machine (SVM) classification [26]. The theory behind the HOG features description is based on local appearance and shape of all objects in an image, which can be described by the distribution of intensity gradients or edge directions. The implementation divides the image into small-connected regions (cells) that can have different shapes (circles or squares). For each cell, a histogram of gradient directions (or edge orientations) for the pixels within the cell is compiled. These histograms are later weighted according to the magnitude of the gradients of the computed cells, and later normalized according to blocks of cells. These blocks of cells can have different shapes and can overlap, thus a given cell can be included in more than one block. The combinations of all these

histograms representing the occurrence of a given angle in each cell inside all the blocks of a certain image represent the descriptor of the image. SVM is used to perform the final classification. One of the main drawbacks of the HOG features vector is the high computational cost required for the computation of the features. Modern high efficient techniques using tools such as parallel processing and the reduction of the image to process (only ROIs provided by laser scanner are processed) allow having real time detection.

#### 4.3. Fusion Algorithm

The Kalman Filter was considered a robust and reliable choice for tracking pedestrians, thanks to the fast acquisition frequency of the sensor. In [27] a model for a Kalman Filter to track pedestrians using the constant velocity model is given, modeling accelerations as system errors. Equations (5) and (6) present the system error  $Q$  and the measurement error covariance matrixes  $R$  of the KF for the model used, which models the variations of the velocity of the pedestrian according to the maximum acceleration amplitude ( $a_x, a_y$ ):

$$Q = \begin{bmatrix} \frac{a_x^2 t^3}{3} & \frac{a_x^2 t^2}{2} & 0 & 0 \\ \frac{a_x^2 t^2}{2} & a_x^2 & \frac{a_y^2 t^3}{3} & \frac{a_y^2 t^2}{2} \\ 0 & 0 & \frac{a_y^2 t^2}{2} & a_y^2 \\ 0 & 0 & a_y^2 & a_y^2 \end{bmatrix} \quad (5)$$

$$R = \begin{pmatrix} \sigma_x^2 & 0 \\ 0 & \sigma_y^2 \end{pmatrix} \quad (6)$$

where  $\sigma_x^2$  y  $\sigma_y^2$  is the standard deviation for the measurements in  $x, y$  coordinates. These deviations have been calculated using test sequences (as will be explained in Section 7). As both systems share the ROI coordinates, the deviation in the measurements are considered equal for both detections. The values  $a_x$  and  $a_y$  in Equation (6) are the maximum amplitude of the acceleration in each axis. According to [28], these amplitudes can be defined as 11 m/s<sup>2</sup> each.

The constant velocity model for the Kalman filter is defined in Equations (7–10):

$$X = \begin{bmatrix} x \\ y \\ v_x \\ v_y \end{bmatrix} \quad (7)$$

where  $X$  is the state vector,  $x, y$  represent the location of the pedestrian (in meters) in relation to the vehicle,  $v_x$  and  $v_y$  are the speed of the pedestrian in meters per second:

$$Y = \begin{bmatrix} x \\ y \end{bmatrix} \quad (8)$$

$Y$  is the measurements vector for the Kalman Filter, with  $x, y$  the location of the new detections,  $H$  is the observation model matrix and  $F$  is the state transition model matrix.

$$H = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} \quad (9)$$

$$F = \begin{bmatrix} 1 & 0 & t & 0 \\ 0 & 1 & 0 & t \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (10)$$

The aforementioned model is used for movement estimation. Several works proved the usability of the constant velocity model to model the movement of pedestrians [28]. Furthermore, the high acquisition rate of the laser scanner (approx. 20 frames per second) which was used as time reference allows a fast adaptation to any speed changes in the movement of the pedestrian.

Data association is the process of combining the new detection with the already existing tracks. The first step is to reduce the possible combination to those obstacle detections located close to the tracks. To do this, a square gate was defined:

$$K_{Gl}\sigma_r \quad (11)$$

where  $\sigma_r$  is the residual standard deviation and  $K_{Gl}$  is a constant that was empirically chosen.

After that, association is performed among those detections that are included within the gate of each track, using normalized distance and a stability factor, giving less priority to less stable measurements:

$$d^2 = \frac{(x_i - \bar{x})^2}{\sigma_x^2} + \frac{(y_i - \bar{y})^2}{\sigma_y^2} + \ln(\sigma_x \sigma_y) \quad (12)$$

## Track Management

For track management, the pedestrian definition follows Equation (13):

$$Pedestrian = \begin{bmatrix} X \\ lp \\ cp \end{bmatrix} \quad (13)$$

where  $X$  is the KF state vector defined in Equation (7) and  $lp$  (laser positive) and  $cp$  (camera positive) represent the Boolean values that indicate if a pedestrian has been positively detected (1 value) by the corresponding sensor over time. Following this definition, each detection was also defined by Equation (14):

$$Detection = \begin{bmatrix} Y \\ New\_lp \\ New\_cp \end{bmatrix} \quad (14)$$

As in Equation (13),  $New\_lp$  and  $New\_cp$  provides a positive value (1) or not (0) of the given sensor for a given detection.

Thus after KF update, the Pedestrian state vector is updated with the state vector of the KF and the values of the laser and camera detection are updated with the Boolean sum (OR) of the corresponding values with the detection values in Equation (15):

$$Pedestrian_k = \begin{bmatrix} \hat{X}_{k|k} \\ (lp_{k-1} + New\_lp_k) \\ (cp_{k-1} + New\_cp_k) \end{bmatrix} \quad (15)$$

where the state of a given pedestrian in an instant  $k$  is defined by the a posteriori state vector provided by the KF  $\hat{X}_{(k|k)}$  and the Boolean values that indicate a positive detection from the creation of the track by any subsystem.

A track is thus considered consolidated with certainty enough to be reported as a pedestrian, when both *cp* and *lp* provide positive values. The logic followed to track creation and deletion is depicted by the following table:

Track creation	-	A track is created if a detection does not match with any track.
	-	Every new track is considered consolidated when both sensors detect it.
Track deletion	-	<i>Not consolidated</i> . If there is no match after three consecutive scans, the track is deleted.
	-	<i>Consolidated</i> . After five no updates.
Track update	-	Track matching with any of the sensors.

It should be pointed out that Assignment Matrix was used to track association, following at least overall cost assignment [28,29].

## 5. V2V Pedestrian Detection Communication

Once the pedestrians are detected, it is important to alert vehicles in the surroundings, using a fast and reliable V2V communication system. This means that the information about the location of pedestrians and the corresponding warnings alerts have to be sent to other vehicles. Two elements have to be defined from the viewpoint of communications. On the one hand, the communication technology used to support the data transmission that must be adapted to vehicular environments; on the other hand, the structure of the data packages that contain the information to be transmitted from one vehicle to the others. In any case, we consider that each vehicle equips a vehicular communication module and can be considered as a node of the network.

### 5.1. Communication Technology

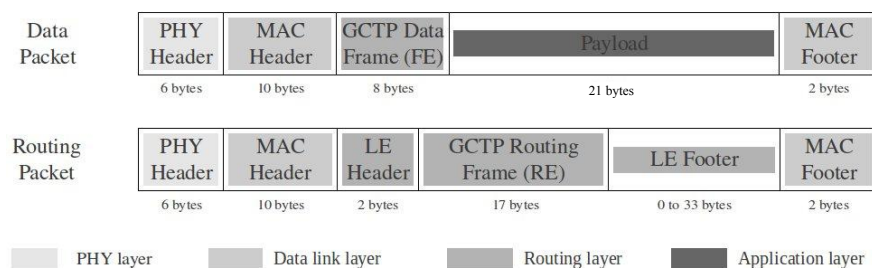
The communication technology used in the experiments related to this paper is based on Maxfor Inc. MTM-CM3100 gateways, based on the TelosB platform. This allows full connectivity with any PC through a USB port, and is used as interface to access a vehicular mesh network. This device works under the TinyOS open code operating system. In order to access the wireless network to 2.4 GHz in mesh, it uses the IEEE 802.15.4 standard at physical and link level and a mesh geo-networking protocol, which guarantees the desired functionality of the Vehicular Ad-Hoc Network (VANE). This GeoNetworking is supported by a novel GeoNetworking protocol (Geocast Collection Tree Protocol—GCTP) implemented to support the mesh routing. The GeoNetworking protocol is a network protocol that resides in the network and transport layer [30] and is executed in each ad hoc router, specifically in each GeoAdhoc router. This novel protocol reconfigures the network structure in accordance with the GPS positions of the network nodes. With this technology it is possible to develop the necessary algorithm to optimize the message routing in V2V and V2I communications. This algorithm can be used as a solution to be applied to other specifically designed technologies when available.

## 5.2. Data Package Structure

The data package structure contains the format of the information to be transmitted from one communication node to another. It contains the necessary data to support the application of V2V-based pedestrian detection warning.

Figure 6 shows the structure of both data packets defined to support the application and the GeoNetworking. Thus, the geo-network routing definition uses the routing packets that allow establishing the necessary route tables in mobility to guarantee an efficient information transmission among the nodes of the vehicular mesh network. In this packet the PHY Header field, MAC header, LE Header, LE Footer and MAC Footer represent the standard information of an IEEE 802.15.4 network [31]. The GCTP Routing Frame represents the geo-referenced information used to provide the GCTP algorithm with the necessary information to define the geo-based route tables of the vehicular mesh network.

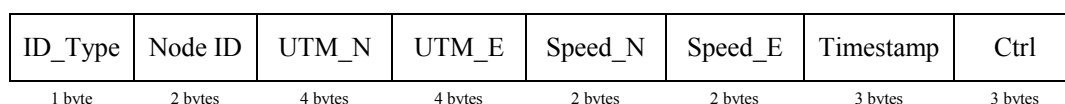
**Figure 6.** Structure of the data packages to support GeoNetworking and V2V pedestrian detection warning.



The data packet schema defines the packet structure to support the application data exchange. Similarly the Routing Packets and the Data Packets are formed by 3 IEEE 802.15.4 standard fields (PHY Header, MAC Header and MAC Footer), a specific Geo-Networking field (GCTP Data Frame) and the Payload that contains the application information.

This Payload is described in detail in Figure 7. The different fields that are needed to support the V2V-based pedestrian warning system are shown. The field ID\_Type identifies the type of obstacle that is detected by the system, which, in the present application, always corresponds to a pedestrian. The Node ID represents the unique identifier associated with each node of the network. UTM\_N and UTM\_E represent the GPS Universal Transverse Mercator (UTM) coordinate for the given pedestrian. Speed\_N and Speed\_E represent the velocity vector of the detected pedestrians, thus all the information related to the status of a given pedestrian, obtained from the KF state, is provided in the payload field. Timestamp is a temporal identifier that represents the instant when the message was generated. Finally, the Ctrl field is a CRC to guarantee the right payload definition.

**Figure 7.** Detail of the Payload of the data packets.



It is important to remark that to allow real time detection, periodic updating of the detections is mandatory. On the other hand, it is also important to reduce the load of the network. To allow a compromise between both constraints, the updating period of the detection was defined as a parameter in both emitter and receiver, allowing the system to be adapted to reduce the workload without losing real time performance.

### 5.3. Data Transmission Operation

The nodes of the network link one to another at a maximum distance of 100 m. The link means that the GeoNetwork is established but no data transmission is required. The data transmission begins in the moment that the pedestrian detection system detects a pedestrian on the route of the vehicle. In this moment, the system starts to broadcast the pedestrian warning through V2V communications. The receptor vehicles retrieve the data frames and their information and decide whether to warn the driver, depending whether or not they are in the influence area of the receiver.

## 6. Test & Results

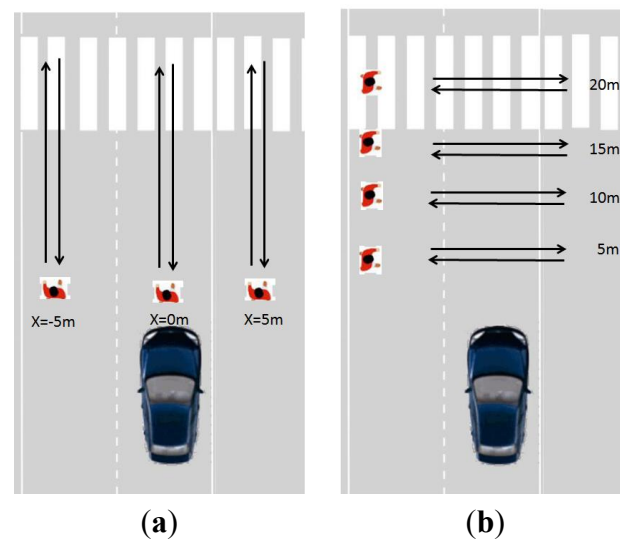
Several tests were performed, divided into three sets. First, a calibration test was performed to tune up the KF (Equation (6)) and to check the accuracy of the obstacle detection algorithm. In the second test, validation tests were performed, where the entire algorithm was tested in a controlled environment obtaining matching results for each sensor performance independently and for the complete system. In this tests set, performance of the system was checked, as well as the parameters of the application, which was varied to obtain the best configuration (e.g., number of negative detections to eliminate a track) to find the final configuration. Finally, tests in real road situations were performed where false positives are more common due to the variety of pedestrians and other unpredicted obstacles.

### 6.1. Calibration Test

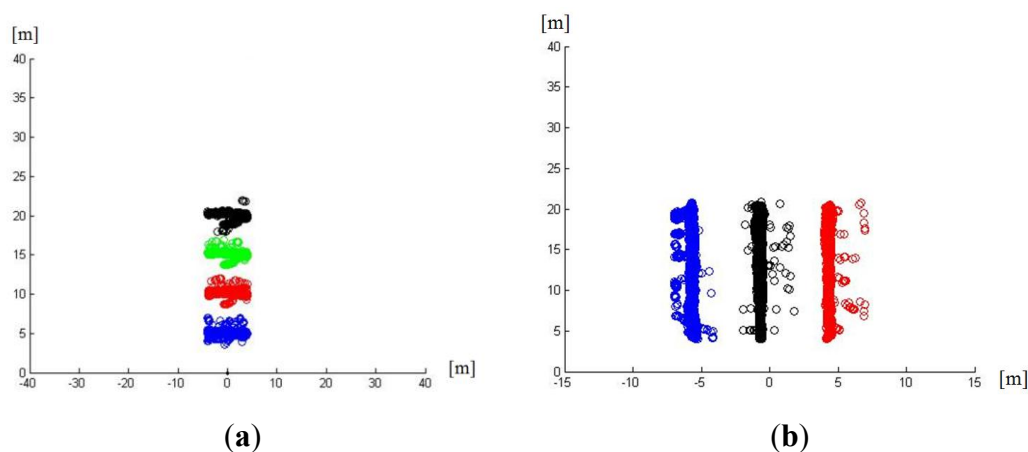
This test consisted of test sequences with a single pedestrian performing lateral and vertical movements. In lateral movements the y coordinate was fixed, so the system could measure the error in the y coordinate as the pedestrian moves along the x axes. For vertical movements, the pedestrian had the x coordinate fixed and moved along the y axis, thus the deviation in x was measured (Figure 8). Furthermore, this test helped to check the performance of the laser scanner detection and segmentation algorithm and the accuracy to the estimation of the movement of the pedestrian. The tests were performed over 20 pedestrians, all performing the same movements, in controlled scenarios. The movements were predefined, providing ground truth, used to check the accuracy of both laser scanner detection and KF estimation. Subsequent tests performed in real road scenarios were used to check the reliability of the detection algorithm.

The results of these test sequences (presented in Figure 9) showed that the assumption of two independent errors for both measurements (x,y) was correct since the error pattern is similar among the lateral movements and vertical movements—regardless of the values on the other axis. Only lateral movements present a higher error when the pedestrian is in the center that was negligible for the present application.

**Figure 8.** The experiments aimed at measuring the measurement errors. (a) x position is fixed. (b) y position is fixed.



**Figure 9.** Results of the experiments to measure the measurement errors y (in meters) vs. x (in meters). (a) Total measurement of the experiments for lateral movement. (b) total measurement of the experiments for vertical movement. (c) y error vs. x coordinates in lateral experiments. (d) x error vs. y coordinate in vertical movement experiments.



Finally, static axis, allowed checking accuracy of the obstacle detection and tracking, providing a standard deviation of 0.15 m in each axis for laser scanner localization and 0.2 m of mean error for KF estimation. This accuracy allows a trustable detection system able to provide obstacle location with high precision.

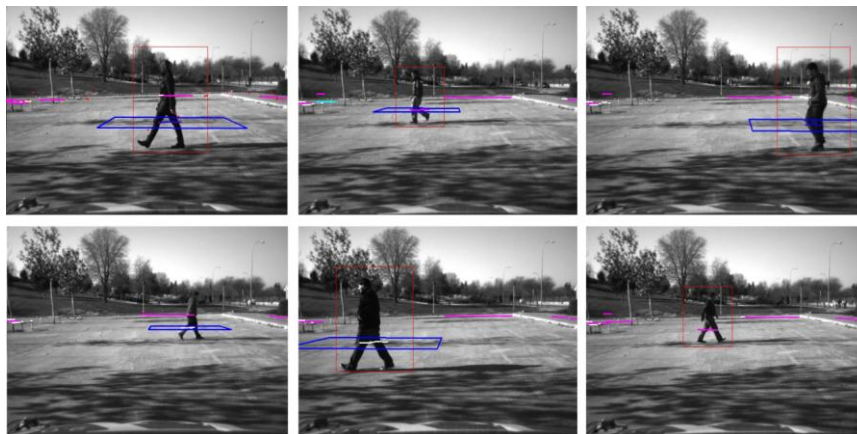
## 6.2. Validation Tests

Validation tests were performed in a parking lot with a single pedestrian simulating pedestrian crossing (Figure 10). Up to nine different pedestrians were tested (Figure 10) when the vehicle was both in movement and stopped. Detections results were 86.2% for computer vision approach, 90.63% for pedestrian approaches and 96.02% for the Fusion system. Here some details need to be remarked on to clarify those results.



- The tests were performed using the best scenario possible. False positives were at a minimum due to the fact that the parking lot represented a controlled scenario where pedestrians were easy to detect, thanks to the absence of other obstacles that could make the detections difficult because of the similarity to pedestrian patterns or occlusions. So false positives found in this test were negligible.
- Laser scanner detection has a higher rate than vision scanner. That could lead to the wrong assumption that the system developed for laser scanner detection is more robust and trustable. As explained in the laser scanner detection subsystem, the limited information provided by the sensor makes classification a challenge, as false positives are common, thereby decreasing the reliability of the detection system. For the present test, as the controlled scenario presented no other obstacles, no false positives were found but it should be taken into account in real scenario tests.
- It is also proved that the detection performance is clearly increased by the fusion approach presented in this paper.

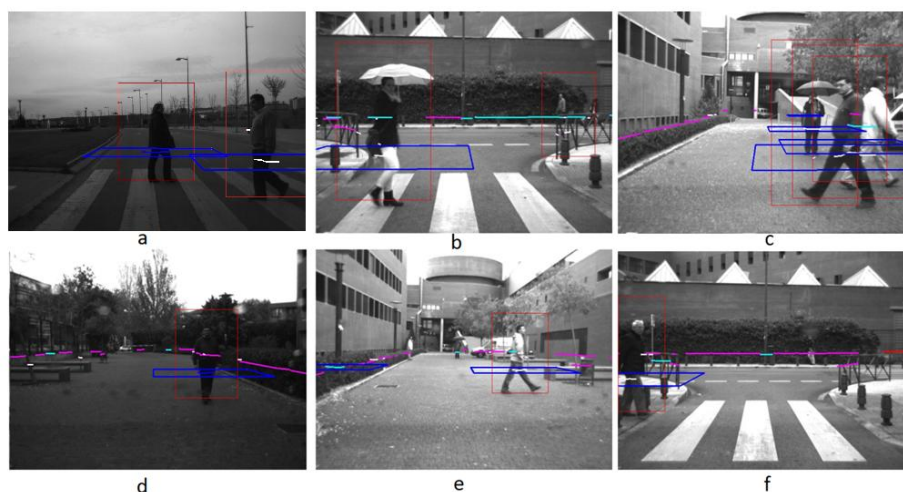
**Figure 10.** Several validation tests results. Red squares stand for positive detections of the camera. Blue squares constitute laser scanner detections. In the image polyline projections in the image space are also represented.



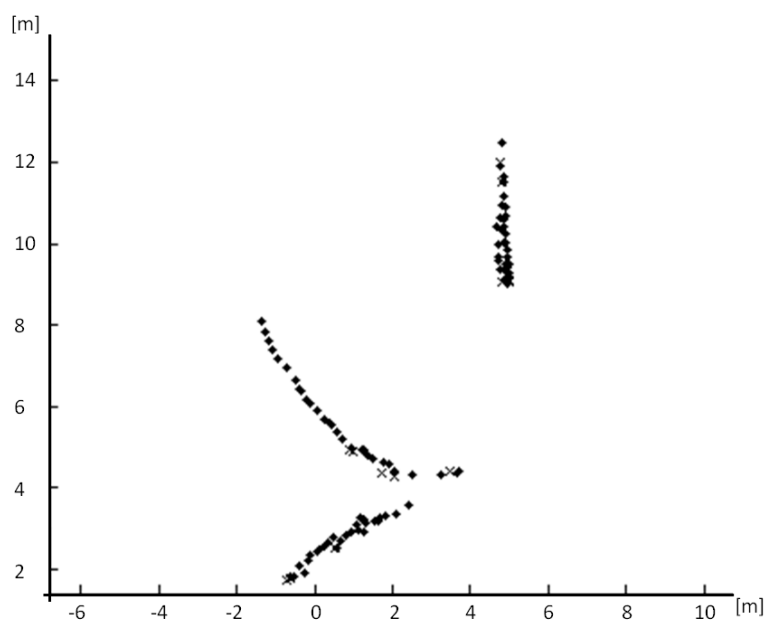
### 6.3. Real Situation Tests

These tests included several pedestrians walking in real road situations including a single pedestrian crossing (Figure 11b, d, e and f) and several pedestrians in crossing trajectories (Figure 11a, c). Tests performed in real roads include both urban and interurban scenarios. The latter represents easier scenarios for pedestrian detections regarding to the absence of other obstacles that could lead to misdetection. Urban scenarios represent more challenging scenarios with higher diversity of obstacles, thus misdetections are expected, although the velocities of the vehicles are lower. Figure 12 shows, as example, the entire tracking for Figure 11c, with several pedestrians involved, in different directions, including a static pedestrian. Here, the algorithm is able to track all the pedestrians even *in situ* where the pedestrians are crossing, with several occlusions. It is also noticeable that the standing pedestrian is detected and tracked from the beginning of the sequence. The high performance of the system is proved in the numerous sequences tested, as depicted in Table 1.

**Figure 11.** Several detections with subsystems detection highlighted. (a) and (b) show two detections. (c) shows three pedestrians crossing. (d), (e) and (f) show single positive detection.



**Figure 12.** Results for the tracking of the pedestrians in Figure 11c, in which three pedestrians are tracked. Dots represent normal detections. Crosses represent detections where the track is not updated. The vehicle was moving at 20 km/h and stops before the pedestrian.



**Table 1.** Results of the overall system with low level comparison.

	Camera		Laser Scanner		Fusion	
	% Positive	% False Positives	% Positive	% False Positives	% Positive	% False Positives
Test	86,2	5.19	90,63	16.23	<b>96.02</b>	<b>0.89</b>
Real Road	71.57	7.39	89.94	16,84	<b>92.85</b>	<b>4.24</b>

In the scope of these tests, the real time performance of the system was tested. The results showed that, by means of the parallelization techniques, modern processors, and by reducing the regions on the images where the visual algorithm searches, real time performance is achieved, adding no delays in the detections. Thus the response time is mainly limited by the acquisition frequency of the sensors. As commented before, the high acquisition frequency of the laser scanner allows it to detect pedestrians with a rate of approximately 20 Hz (56 milliseconds). At high speed (120 km/h) it represents less than 2 m covered by the vehicle in each scan, providing real time performance. Thus the limitations of the system presented are linked to the limitations inherent to the sensors, but not related to the work presented here (*i.e.*, the laser scanner detection distance, or computer camera field of view).

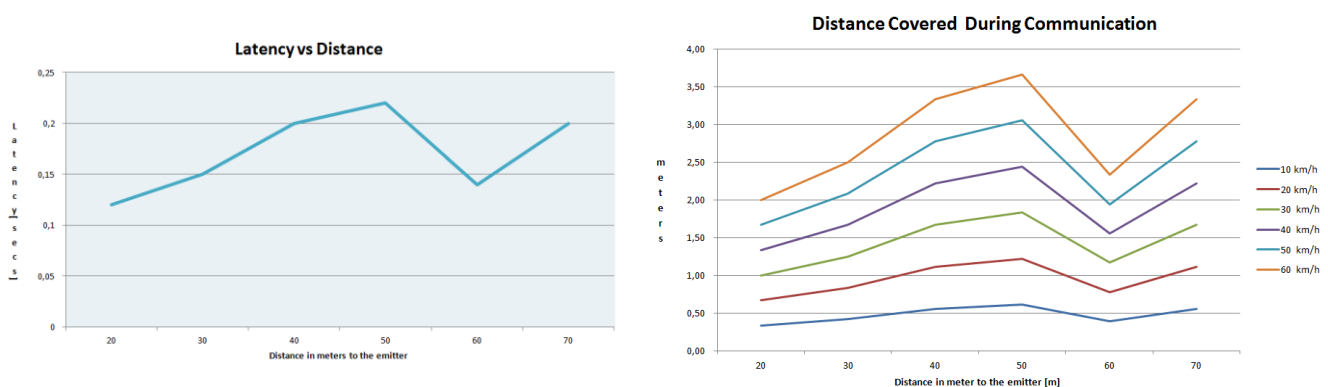
#### 6.4. Communication Tests

After checking the performance of the detection algorithm, the network performance was tested to prove the viability of a VANET to alert vehicles in the surroundings, when a pedestrian is detected. The test to measure the latency in the communication process was performed. This test checked the time necessary to send the necessary data to alert another vehicle of a given detection. The performance of the network guarantees the effectiveness of the assistance system, maintaining the average latency lower than 0.17 s, proved in a set of tests at several distances in the operating range. The average behavior of the network is demonstrated in Figure 13.

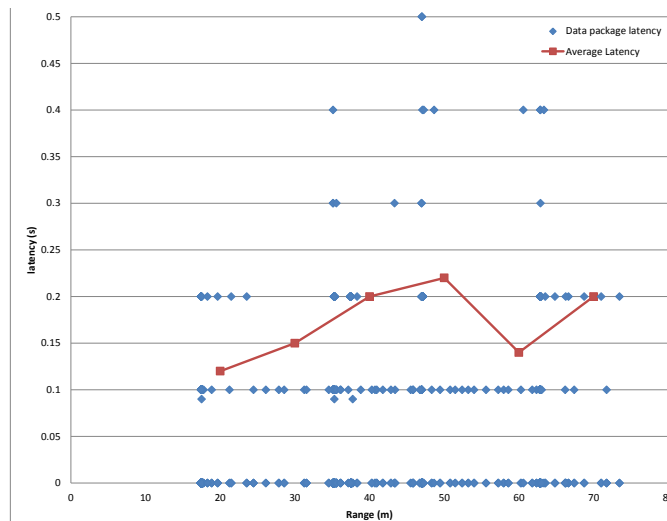
Finally, in Figure 14, the performance of the network in these latency tests is shown graphically, combining several distances, at a data packets transmission rate of 4 Hz during 384 s, obtaining a data loss rate of 0.2%.

This test proved that the performance of the communication protocol is very efficient and effective for safety application with a loss rate of approximately 0.2%, proving the high reliability of the system. Besides, the test performed showed a fast response able to provide alarm communication within a short distance covered by the vehicle, as shown in Figure 13.

**Figure 13.** Network performance. Left, latency results vs. distance to the emitter. Distance covered by the receiver, according to the velocity and the distance to the emitter (given the latency of the communication).



**Figure 14.** Results of the performance tests of the mesh GeoNetwork that supports the assistance system.



## 7. Conclusions

The paper presented provides a step forward in the following aspects:

- An improved laser scanner approach that provides pedestrian detection with the limited information provided by the laser scanner, by means of a pattern matching algorithm augmented with tracking stage and context information.
- Fusion algorithm, able to combine detections from the laser scanner and computer vision approach at a high level.
- An efficient and reliable VANET protocol able to provide information of the detected pedestrians to the vehicles in the surroundings with time enough in advance.

Tests under real conditions proved the viability of the algorithms and the performance of the system in challenging scenarios. The viability of the laser scanner for pedestrian detection was proved (up to 92%). Although a lot of work must still be performed in order to reduce the amount of false positives to values suitable for a single sensor safety application, the laser scanner overall performance was proved to be very useful when combined with the camera detections.

The novel fusion approach for pedestrian detection has proved that, by adding laser scanner pedestrian detection and context information, performance and trustability can be increased, fulfilling the demanding requirements required for traffic safety applications. Furthermore the tracking stage presented was able to track the different pedestrians in variable urban scenarios with high accuracy and trustability.

On the other hand, the efficient V2V communication protocol proved to be able to alert surrounding vehicles even at distances up to 100 m with time in advance to warn the drivers of the dangerous situation, allowing avoiding accidents caused by pedestrian occlusions. The present work has proved that, by means of an effective data fusion algorithm, and a communication protocol, it is possible to create a powerful tool able to increase road safety. This represents a step forward, benefiting from the advantages of the state-of-the-art sensors available in order to provide a novel application able to give multimodal detection.

## Acknowledgments

This work was supported by the Spanish Government through the Cicyt projects (GRANT TRA2010-20225-C03-01, GRANT TRA2010-20225-C03-03, GRANT TRA 2011-29454-C03-02 and iVANET TRA2010-15645) and CAM through SEGVAUTO-II (S2009/DPI-1509).

## Conflicts of Interest

The authors declare no conflict of interest.

## References

1. European Commission. Mobility and Transport Statistics. Available online: [http://ec.europa.eu/transport/road\\_safety/index\\_en.htm](http://ec.europa.eu/transport/road_safety/index_en.htm) (accessed on 1 July 2013).
2. Jiménez, F.; Naranjo, J.; García, F. An improved method to calculate the time-to-collision of two vehicles. *Int. J. Intell. Transp. Syst. Res.* **2013**, *11*, 34–42.
3. Jiménez, F.; Naranjo, J.E.; Gómez, O. Autonomous manoeuvring systems for collision avoidance on single carriageway roads. *Sensors* **2012**, *12*, 16498–16521.
4. Garcia, F.; de la Escalera, A.; Armingol, J.M.; Herrero, J.G.; Llinas, J. Fusion Based Safety Application for Pedestrian Detection with Danger Estimation. In Proceedings of the 14th International Conference on Information Fusion (FUSION), Chicago, IL, USA, 5–8 July 2011; pp. 1–8.
5. Garcia, F.; de la Escalera, A.; Armingol, J.M.; Jimenez, F. Context Aided Fusion Procedure for Road Safety Application. In Proceedings of 2012 IEEE Conference on Multisensor Fusion and Integration for Intelligent Systems (MFI), Hamburg, Germany, 13–15 September 2012; pp. 407–412.
6. García, F.; Jiménez, F.; Naranjo, J.E.; Zato, J.G.; Aparicio, F.; Armingol, J.M.; de la Escalera, A. Environment perception based on LIDAR sensors for real road applications. *Robotica* **2012**, *30*, 185–193.
7. Gerónimo, D.; Sappa, A.D.; Ponsa, D.; López, A.M. 2D-3D-based on-board pedestrian detection system. *Comput. Vis. Image Underst.* **2010**, *114*, 583–595.
8. Geronimo, D.; Lopez, A.M.; Sappa, A.D.; Graf, T. Survey of pedestrian detection for advanced driver assistance systems. *IEEE Trans. Pattern Anal. Mach. Intell.* **2010**, *32*, 1239–1258.
9. Llorca, D.F.; Sotelo, M.A.; Hellín, A.M.; Orellana, A.; Gavilan, M.; Daza, I.G.; Lorente, A.G. Stereo regions-of-interest selection for pedestrian protection: A survey. *Transp. Res. Part C Emerg. Technol.* **2012**, *25*, 226–237.
10. Milanés, V.; Llorca, D.F.; Villagrà, J.; Pérez, J.; Parra, I.; González, C.; Sotelo, M.A. Vision-based active safety system for automatic stopping. *Expert Sys. Appl.* **2012**, *39*, 11234–11242.
11. Premebida, C.; Monteiro, G.; Nunes, U.; Peixoto, P. A Lidar and Vision-Based Approach for Pedestrian and Vehicle Detection and Tracking. In Proceedings of IEEE Intelligent Transportation Systems Conference ITSC, Seattle, WA, USA, 30 September–3 October 2007; pp. 1044–1049.
12. Spinello, L.; Siegwart, R. Human Detection Using Multimodal and Multidimensional Features. In Proceedings of 2008 IEEE International Conference on Robotics and Automation, Pasadena, CA, USA, 19–23, May 2008; pp. 3264–3269.

13. Premebida, C.; Ludwig, O.; Nunes, U. LIDAR and vision-based pedestrian detection system. *J. Field Robotics* **2009**, *26*, 696–711.
14. Premebida, C.; Ludwig, O.; Silva, M.; Nunes, U. A Cascade Classifier Applied in Pedestrian Detection Using Laser and Image-Based Features. In Proceedings of IEEE Intelligent Transportation Systems Conference ITSC, Madeira Island, Portugal, 19–22 Sept. 2010; pp. 1153–1159.
15. Kaempchen, N.; Buehler, M.; Dietmayer, K. Feature-Level Fusion for Free-Form Object Tracking Using Laserscanner and Video. In Proceedings of IEEE Intelligent Vehicles Symposium 2005, Las Vegas, ND, USA, 6–8 June 2005; pp. 453–458.
16. Hwang, J.P.; Cho, S.E.; Ryu, K.J.; Park, S.; Kim, E. Multi-Classifer Based LIDAR and Camera Fusion. In Proceedings of IEEE Intelligent Transportation Systems Conference ITSC, Seattle, WA, USA, 30 September–3 October 2007; pp. 467–472.
17. Szarvas, M.; Sakai, U. Real-time Pedestrian Detection Using LIDAR and Convolutional Neural Networks. In Proceedings of 2006 IEEE Intelligent Vehicles Symposium, Tokyo, Japan, 13–15 June 2006; pp. 213–218.
18. Ludwig, O.; Premebida, C.; Nunes, U.; Ara, R. Evaluation of Boosting-SVM and SRM-SVM Cascade Classifiers in Laser and Vision-Based Pedestrian Detection. In Proceedings of IEEE Intelligent Transportation Systems Conference ITSC, Washington, DC, USA, 5–7 October 2011; pp. 1574–1579.
19. Pérez, G.A.; Frolov, V.; Puente León, F. Information fusion to detect and classify pedestrians using invariant features. *Inform. Fusion* **2010**, *12*, 284–292.
20. Broggi, A.; Cerri, P.; Ghidoni, S.; Grisleri, P.; Jung, H.G. Localization and Analysis of Critical Areas in Urban Scenarios. In Proceedings of IEEE Intelligent Vehicles Symposium, Beijing, China, 12–15 October 2008; pp. 1074–1079.
21. Misener, J.A. *VII California: Development and Deployment Proof of Concept and Group-Enabled Mobility and Safety (GEMS)*; California PATH Research Report, UCB-ITS-PRR-2010-26; Department of Transportation, University of California: Berkeley, CA, USA, 2010.
22. Schnacke, D. Applications Supported by Next-Generation 5.9 GHz DSRC. In Proceedings of ITS World Congress, Madrid, Spain, 18 November 2003.
23. Glaser, S.; Nouveliere, L.; Lusetti, B. Speed Limitation Based on an Advanced Curve Warning System. In Proceedings of 2007 IEEE Intelligent Vehicles Symposium, Seattle, WA, USA, 30 September–3 October 2007; pp. 686–691.
24. Buchenscheit, A.; Schaub, F.; Kargl, F.; Weber, M. A VANET-Based Emergency Vehicle Warning System. In Proceedings of 2009 IEEE Vehicular Networking Conference VNC, Tokyo, Japan, 28–30 October 2009; pp. 1–8.
25. Garcia, F.; Musleh, B.; de Escalera, A.; Armingol, J.M. Fusion Procedure for Pedestrian Detection Based on Laser Scanner and Computer Vision. In Proceedings of 2011 14th International IEEE Conference on Intelligent Transportation Systems (ITSC), Washington, DC, USA, 5–7 October 2011; pp. 1325–1330.
26. Dalal, N.; Triggs, B. Histograms of Oriented Gradients for Human Detection. In Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2005 (CVPR 2005), San Diego, CA, USA, 20–26 June 2005; Volume 1, pp. 886–893.

27. Kohler, M. *Using the Kalman Filter to track Human Interactive Motion—Modelling and Initialization of the Kalman Filter for Translational Motion*; Technical Report 629; Informatik VII, University of Dortmund: Dortmund, Germany, 1997.
28. Blackman, S.S. *Multiple-Target Tracking with Radar Application*; Artech House: Norwood, MA, USA, 1986.
29. Blackman, S.; Popoli, R. *Design and Analysis of Modern Tracking Systems*; Artech House: Norwood, MA, USA, 1999.
30. *European Standard (Telecommunications Series): Intelligent Transport Systems (ITS); Communications Architecture*; ETSI EN 302 665 (2010-09); ETSI: Sophia Antipolis, France, 2010.
31. Gnawali, O.; Fonseca, R.; Jamieson, K.; Levis, P. *CTP: Robust and Efficient Collection through Control and Data Plane Integration*; Technical Report SING-08-02; University of Southern California: Los Angeles, CA, USA, 2008.

© 2013 by the authors; licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/3.0/>).